

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.

MENU

SEARCH

INDEX

DETAIL

1/1



JAPANESE PATENT OFFICE

## PATENT ABSTRACTS OF JAPAN

(11)Publication number: 07129593

(43)Date of publication of application: 19.05.1995

(51)Int.Cl.

G06F 17/27  
G06F 17/30

(21)Application number: 06061885

(71)Applicant:

TOSHIBA CORP

(22)Date of filing: 31.03.1994

(72)Inventor:

SHIMOGOORI NOBUHIRO

(30)Priority

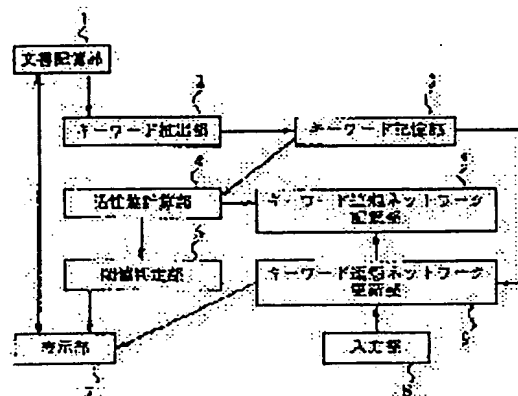
Priority number: 05222924 Priority date: 08.09.1993 Priority country: JP

(54) TEXT SELECTING DEVICE

(57)Abstract:

PURPOSE: To provide a text selecting device capable of selecting and providing a text which is supposed to be valuable for user.

CONSTITUTION: After extracting a key word from the objective text by a key word extracting part 2 and calculating a node active value on the key word associative network of the key word by an active value calculation part 4, the device propagates the active value onto the key word associative network, calculates the score of the objective text based on the active value of all the nodes on the key word associative network after active propagation and information on weight and displays the text on a display part 7 when the calculated score is not less than a prescribed threshold value.



LEGAL STATUS

[Date of request for examination]  
[Date of sending the examiner's decision of rejection]  
[Kind of final disposal of application other than the  
examiner's decision of rejection or application converted  
registration]  
[Date of final disposal for application]  
[Patent number]  
[Date of registration]  
[Number of appeal against examiner's decision of rejection]  
[Date of requesting appeal against examiner's decision of  
rejection]  
[Date of extinction of right]

Copyright (C); 1998 Japanese Patent Office

MENU

SEARCH

INDEX

DETAIL

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

K 0 6 F - 3 2

(11) 特許出願公開番号

特開平7-129593

(43) 公開日 平成7年(1995)5月19日

(51) Int.Cl.<sup>6</sup>

G 0 6 F 17/27

17/30

識別記号

庁内整理番号

F I

技術表示箇所

7315-5L

9194-5L

G 0 6 F 15/ 20

15/ 403

5 5 0 F

3 2 0 Z

審査請求 未請求 請求項の数7 O L (全 26 頁)

(21) 出願番号 特願平6-61885

(22) 出願日 平成6年(1994)3月31日

(31) 優先権主張番号 特願平5-222924

(32) 優先日 平5(1993)9月8日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 下郡 信宏

神奈川県川崎市幸区柳町70番地 株式会社

東芝柳町工場内

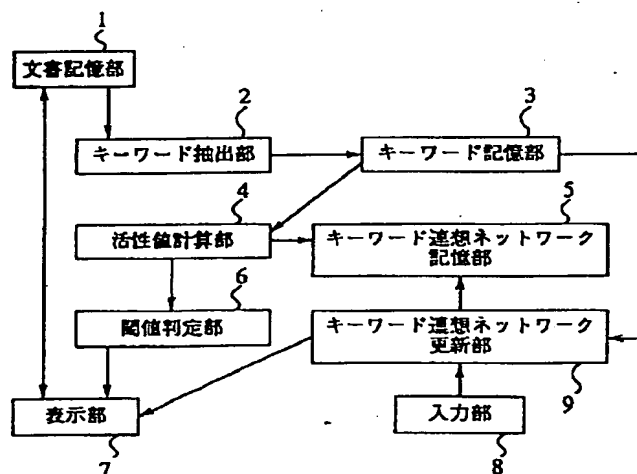
(74) 代理人 弁理士 鈴江 武彦

(54) 【発明の名称】 テキスト選定装置

(57) 【要約】

【目的】 ユーザにとって価値あると思われるテキストを選定して提示するテキスト選定装置を提供する。

【構成】 対象としているテキストからキーワード抽出部2でキーワードを抽出し、このキーワードのキーワード連想ネットワーク上におけるノード活性値を活性値算出部4で算出した後に、この活性値をキーワード連想ネットワーク上に伝搬させ、活性伝搬後におけるキーワード連想ネットワーク上の全ノードの活性値と重みに関する情報とに基づいて対象としているテキストの得点を算出し、算出された得点が所定の閾値以上のときに該テキストを表示部7に表示するように構成されている。



## 【特許請求の範囲】

【請求項1】個々のキーワードに対応した大きさに関する情報および重みに関する情報を持つノードおよびノード間を接続するリンクを備えたキーワード連想ネットワークを用いてユーザが所望としているテキストの選定を行うテキスト選定装置において、

前記キーワード連想ネットワーク上の全ノードの活性値を初期状態に設定する手段と、

対象としているテキストから所定のキーワードを抽出するキーワード抽出手段と、

この手段により抽出されたキーワードが前記キーワード連想ネットワーク上にノードとして存在するか否かを判断する判断手段と、

この判断手段によって前記抽出されたキーワードがノードとして存在すると判断されたときに前記大きさに関する情報に基づいて上記抽出されたキーワードに対応するノードの活性値を算出する活性値算出手段と、

この手段によって算出されたノードの活性値を前記キーワード連想ネットワーク上に伝搬させる活性伝播手段と、

活性伝搬後における前記キーワード連想ネットワーク上の全ノードの活性値と重みに関する情報とに基づいて前記対象としているテキストの得点を算出する得点算出手段と、

この手段により算出されたテキストの得点が所定の閾値以上のときに該テキストをユーザに提示する提示手段とを備えてなることを特徴とするテキスト選定装置。

【請求項2】前記活性値算出手段は、前記テキストから抽出されたキーワードについて、前記キーワード連想ネットワーク上の各キーワードを中心とする所定領域の部分ネットワークについて活性値を求めるものであることを特徴とする請求項1に記載のテキスト選定装置。

【請求項3】前記部分ネットワークは、各キーワードを中心とする前記キーワード連想ネットワーク上の距離によって決められたものであることを特徴とする請求項2に記載のテキスト選定装置。

【請求項4】前記部分ネットワークは、各キーワードを中心とする前記キーワード連想ネットワーク上のリンクの強さによって決められたものであることを特徴とする請求項2に記載のテキスト選定装置。

【請求項5】前記提示手段は、提示されたテキストに対する評価値を入力する手段と、この手段により入力された評価値に基づいて前記抽出されたキーワードに対応するノードの大きさに関する情報を更新する手段とを含んでいることを特徴とする請求項1に記載のテキスト選定装置。

【請求項7】入力されたテキスト情報の中からユーザに有用と思われるテキスト情報を取出して重要度を予測し、重要と判定されたテキスト情報をユーザに表示提示するテキスト選定装置本体と、

表示されたテキスト情報をユーザが読む際の視線の動きを追跡し、この追跡情報から表示されているテキスト情報に対するユーザの興味の度合を評価して前記テキスト選定装置本体にフィードバックする評価手段と、

この評価手段によって得られた評価値と前記テキスト選定装置本体で用いた予測値とを用いて上記評価手段での評価の際に用いる係数を学習する学習手段とを具備してなることを特徴とするテキスト選定装置。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、テキスト選定装置に係り、特に、新たに与えられたテキストの中からユーザにとって価値があると思われるテキストをたとえばキーワード連想ネットワークを用いて選定するテキスト選定装置に関する。

## 【0002】

【従来の技術】近年、さまざまな情報が氾濫し、文献の累増をもたらしている。これらの文献にはユーザにとって有益なものが多く含まれている。しかし、これらの文献を利用しようとしても、あまりにも文献数が多いため、必要な文献への的確なアクセスが困難になっている。

【0003】そこで、膨大な文献の中からユーザにとって有益なものを選定するために、テキスト検索手段を用いて選定するテキスト選定装置が利用されつつある。このテキスト検索手段は、キーワードによる絞り込みと、キーワード空間内のベクトルを利用するものとの2種類に大別される。

【0004】キーワードによる絞り込みとしては、単にユーザが指定したキーワードをそのまま利用するものや、キーワード間の類似性を定義したシソーラス情報を用いてユーザが指定したキーワードおよび類似したキーワードまでをも含めたキーワード群により選定を行なう方法などが知られている。

【0005】一方、キーワード空間のベクトルとして表現する手法としては、過去に使われたテキスト内のキーワードを予め抽出し、この抽出されたN個のキーワードを用いてN次元のキーワード空間を作成した後、作成したキーワード空間内で該テキストの位置を表現し、これがユーザの興味の範囲内であるか否かを判定することにより選定を行なうものがある(Peter W. Foltz and Susan T. Dumais(PERSONALIZED Information Delivery:ANALYSIS of Information Filtering Methods[COMMUNICATIONS OF THE ACM Dec 1992 Vol135, No.12 pp51-pp60]))。

【0006】また、キーワード空間を動的に扱う方法と

して、類似するキーワード同士を接続して、本来のキーワードのみならず類似するキーワードからもテキストを選定するキーワードコネクションを利用する方法がある(特開平2-42564号公報、特開平2-125363号公報)。かかる文書検索装置では、ユーザが入力したキーワードに基づき、キーワード関連表を用いてデータベース中のテキストの確度を計算し、閾値以上のものを抽出して検索結果としている。

【0007】しかし、キーワードによる絞り込みやキーワード空間のベクトルとして表現する方法を用いてテキストの選定を行った場合には、機械的検索に頼るため、ユーザ特有のキーワード間の類似性を柔軟に表現することができない問題があった。また、重要度の低いキーワードを数多く含むテキストにあっては、キーワードがノイズに埋もれてしまい、抽出することが困難であった。

【0008】さらに、新しいキーワードが生じた場合に際し、個々のキーワードを逐次ユーザが入力しているため、シソーラス辞書やキーワード空間の更新に多大な労力や計算資源等が必要であるという問題もあった。

【0009】また、キーワードコネクションを用いる方法では、ユーザ固有のキーワード間の類似情報を保持しつつ、与えられたキーワードからユーザにとって関連が深いと思われるテキストを検索することは可能であるが、不特定のテキストに対してどの程度興味を持っているかということを判定することが不可能であった。

【0010】たとえば、ある文献についてユーザが自分にとって興味のあるものかどうかを判断する場合には、一般に、ユーザは内容を一つ一つじっくり読んでいるのではなく、文章全体を眺めるように読んでゆき、非常に重要なキーワードが現れたときや、さほど重要でないキーワードでもそれらが数多く現れたときに、さらにじっくりと読むという二重の行動パターンをとることがある。前者のような行動パターンの場合にはキーワードによる従来の検索手法で十分対応できるが、後者のような場合には所望のテキストがノイズに埋もれてしまう可能性がある。

【0011】そこで、このような人間の判定基準を計算機に取り込むため、人間の行動を情報処理装置としてモデル化した認知モデルを用いる試みがなされている。たとえば、John R. Anderson 氏の提唱するACT\* (The Architecture of Cognition: Harvard University Press 1983, ISBN 0-674-04426-6) もこのような認知モデルの一つとして知られている。この理論は、人間の持つ概念をネットワーク上の一つのノードとして表し、このネットワーク上に活性を伝搬させ、活性が高いものが現在思い出している概念であるとするにより、人間の連想記憶能力を表現する。

【0012】しかし、このような認知モデル化された情報処理装置としてテキスト選定装置を構成した場合、データベースの全てのテキスト群に対して全てのキーワー

ド群をキーワード連想ネットワーク上で活性化させて評価しなければならないため、いわゆる計算爆発が起こり非現実的である。

【0013】

【発明が解決しようとする課題】上述のように、従来技術によるテキスト選定は、機械的検索に頼るためユーザ特有のキーワード間の類似性を柔軟に表現することができない問題があった。また、重要度の低いキーワードが数多く出現するテキストの抽出は、キーワードがノイズに埋もれてしまい困難であった。さらに、新しいキーワードが生じた場合に際し、逐次ユーザが入力しているため、シソーラス辞書やキーワード空間の更新に多大の労力や計算資源等が必要であるという問題があった。

【0014】また、キーワードコネクションを用いる方法では、ユーザ固有のキーワード間の類似情報を保持しつつ、与えられたキーワードからユーザにとって関連が深いと思われるテキストを検索することは可能だが、不特定のテキストに対してどの程度興味をもっているかということを判定することが不可能であった。

【0015】さらに、認知モデル化された情報処理装置としてテキスト検索装置を構成した場合、データベースの全てのテキスト群に対して全てのキーワード群をキーワード連想ネットワークで活性化させて評価しなければならないため、計算爆発が起こり非現実的であった。

【0016】そこで、本発明の第1の目的は、ユーザ固有のキーワード間の類似性を柔軟に表現できるテキスト選定装置を提供することである。また、第2の目的は、重要度の低いキーワードが数多く出現するテキストに対しても、テキストの選定を可能とするテキスト選定装置を提供することである。さらに、第3の目的は、キーワード間の類似性情報の管理を容易ならしめるテキスト選定装置を提供することである。また、第4の目的は、キーワード連想ネットワークを用いた際に問題となる計算爆発をなくし、極めて高速にかつ現実的にテキストの選定を行うことができるテキスト選定装置を提供することである。

【0017】

【課題を解決するための手段】上記課題を解決するために、本発明者は、新たなテキストのみを対象にして考えた場合には、キーワード群をキーワード連想ネットワークで作り出すのではなく、テキスト全体を評価し、これを閾値で判定するようにすれば、従来ならばキーワード単位で閾値を設けて切り捨てていたキーワードをも考慮したテキスト選定を行えることに着目した。

【0018】すなわち、請求項1のテキスト選定装置は、個々のキーワードに対応した大きさに関する情報および重みに関する情報を持つノードおよびノード間を接続するリンクを備えたキーワード連想ネットワークを用いてユーザが所望としているテキストの選定を行うテキスト選定装置において、前記キーワード連想ネットワー

10

20

30

40

50

ク上の全ノードの活性値を初期状態に設定する手段と、対象としているテキストから所定のキーワードを抽出するキーワード抽出手段と、この手段により抽出されたキーワードが前記キーワード連想ネットワーク上にノードとして存在するか否かを判断する判断手段と、この判断手段により前記抽出されたキーワードがノードとして存在すると判断されたときに前記大きさに関する情報に基づいて前記抽出されたキーワードに対応するノードの活性値を算出する活性値算出手段と、この手段により算出されたノードの活性値を前記キーワード連想ネットワーク上に伝搬させる活性伝搬手段と、活性伝搬後における前記キーワード連想ネットワーク上の全ノードの活性値と重みに関する情報とに基づいて前記対象としているテキストの得点を算出する得点算出手段と、この手段により算出されたテキストの得点が所定の閾値以上のときに該テキストをユーザに提示する提示手段とを備えたことを特徴としている。

【0019】なお、前記活性値算出手段は、前記テキストから抽出されたキーワードについて、前記キーワード連想ネットワーク上の各キーワードを中心とする所定領域の部分ネットワークについて活性値を求めるものであり、計算量を大幅に低減できるので好ましい。部分ネットワークは、各キーワードを中心とする前記キーワード連想ネットワーク上の距離によって決めてもよいし、各キーワードを中心とする前記キーワード連想ネットワーク上のリンクの強さによって決めてもよい。

【0020】また、前記提示手段は、提示されたテキストに対する評価値を入力する手段と、この手段により入力された評価値に基づいて前記抽出されたキーワードに対応するノードの大きさに関する情報を更新する手段とを含んでいることが好ましい。

【0021】また、前記提示手段は、提示されたテキストをユーザが読む際の視線の動きを追跡し、この追跡情報からテキストの価値を評価する評価手段を備えていることが好ましい。

【0022】請求項7のテキスト選定装置は、入力されたテキスト情報の中からユーザに有用と思われるテキスト情報を取出して重要度を予測し、重要と判定されたテキスト情報をユーザに表示提示するテキスト選定装置本体と、表示されたテキスト情報をユーザが読む際の視線の動きを追跡し、この追跡情報から表示されているテキスト情報に対するユーザの興味の度合を評価して前記テキスト選定装置本体にフィードバックする評価手段と、この評価手段によって得られた評価値と前記テキスト選定装置本体で用いた予測値とを用いて上記評価手段での評価の際に用いる係数を学習する学習手段とを備えている。

【0023】

【作用】請求項1のテキスト選定装置では、まず、新たに対象としているテキストから所定のキーワードを抽出

する。抽出されたキーワードが、キーワード連想ネットワーク上に存在する場合、すなわち各キーワードに対応するノードがキーワード相互間の接続関係を表したキーワード連想ネットワーク上に存在する場合には、そのノードの大きさにより決まる初期値をノード活性値とする。ノードが存在する全てのキーワードについてノードの活性値を求めた後、この活性値をキーワード連想ネットワーク上に後述するACT\*の方法により伝搬させ、最終的なキーワード連想ネットワーク全体のノードの活性値を求める。

【0024】次に、キーワード連想ネットワーク上の全てのノードの活性値とそのノードの重要度を表す重みとを考慮した関数により各々のノードの得点を求め、全てのノードの得点を合計することにより、個々のユーザにとっての価値を表すテキストの得点を求める。

【0025】このようにして求めたテキストの得点がユーザの定める閾値を越えているとき、このテキストをユーザに提示する。ユーザはこのテキストについて判定を行って、その判定結果をシステムに与える。これにより、テキストから抽出された全てのキーワードに対応するノードの大きさをテキストの得点分だけ増加させ、またテキストから抽出されたキーワードに対応するノード相互間にリンクを持たせる。以上のようにして、キーワード連想ネットワークの学習が行なわれる。

【0026】一方、抽出されたキーワードがキーワード連想ネットワーク上にキーワードとして存在しない場合、すなわちキーワードに対応するノードがキーワード連想ネットワーク上に存在しない場合には、上記テキストの選定を行った後、キーワード連想ネットワークについて学習を行なう。すなわち、該ノードをキーワード連想ネットワークに加え、ユーザに該キーワードを表示して重みの入力を促し、この入力値を対応するノードの重みに設定する。同様に、テキストから抽出された全てのキーワードに対応するノードの大きさをテキストの評価値分だけ増加させ、また提示されたテキストから抽出されたキーワードに対応するノード相互間にリンクを持たせることによって、キーワード連想ネットワークの学習を行う。

【0027】したがって、本発明に係るテキスト選定装置によれば、柔軟に新しいキーワードを取り込むことが可能となり、ユーザ固有のキーワード間の類似性情報を得ることができる。

【0028】請求項1のテキスト選定装置において、キーワード連想ネットワーク上で活性を伝搬させる方法としてはAnderson氏の提唱するACT\*論理を用いている。以下に、その活性を伝搬させる方法について説明する。ノード*i*における時間*t*の活性の変化量は、

【0029】

【数1】

7

8

$$\frac{da_i(t)}{dt} = Bn_i(t) - \rho^* a_i(t) \quad \text{--- (1)}$$

ただし、 $a_i(t)$ : ノード  $i$  の時間  $t$  における活性値

$n_i(t)$ : ノード  $i$  の時間  $t$  における総流入活性

$\rho^*$ : 減衰率 (定数)

$B$ : 定数

により表される。

【0030】(1) 式の右辺の第1項はノード  $i$  の時間  $t$  において獲得する活性の総量を表し、第2項はノード  $i$  の時間  $t$  において失われる活性の総量を表している。両\*

\* 者の差分により変化量を求めることができる。したがって、ノード  $i$  の時間  $t$  における活性の総流入量は、

【0031】

【数2】

$$n_i(t) = C_i^*(t) + \sum_j r_{ji} a_j(t - \delta t) \quad \text{--- (2)}$$

ただし、 $C_i^*(t)$ : ノード  $i$  の時間  $t$  においてシステムから供給される活性値の量

により表される。

【0032】(2) 式における右辺の第1項はノード  $i$  の時間  $t$  におけるシステムから直接供給される活性の量を表し、第2項はノード  $i$  の時間  $t$  における他のノードへの活性の流出と流入量を表す。すなわち、活性の流れはこの2種類しかないため、両者の和によりノード  $i$  の時間  $t$  における活性の総流入量  $n_i(t)$  となる。 ※

※ 【0033】ところで、(1) 式は、各ノードに関する式であったが、全てのノードの活性値のベクトル  $A$  と全てのノードの総流入活性のベクトル  $N$  と活性の変化量を表現すると以下ようになる。

【0034】

【数3】

$$\frac{dA_i(t)}{dt} = BN_i(t) - \rho^* A_i(t) \quad \text{--- (3)}$$

同様に、ノード間のリンクの強さを表す行列  $R$  と全ての活性供給量を表す行列  $C$  を導入すると、(2) 式は以下のようにになる。 ★

★ 【0035】

【数4】

$$n_i(t) = C_i^*(t) + RA(t - \delta t) \quad \text{--- (4)}$$

【0036】(3) 式は活性の変化量を求める式であるが、我々が所望としているものは定常状態におけるネットワークの活性の分布を知ることである。定常状態における活性値の変化量は0であるため、(3) 式の左辺は0とみなすことができる。純粋な数学的意味においての定 ☆

☆ 常状態になるためには限りなく長い時間がかかるが、ここでは0と近似することで足りる。したがって、(3) 式および(4) 式から(5) 式を得る。

【0037】

【数5】

$$A = C + \rho RA \quad \text{--- (5)}$$

ただし、 $C = \frac{BC^*}{\rho^*}$

$\rho = \frac{B}{\rho^*}$

この(5) 式を変形することにより最終的に以下の式を得る。 ◆

◆ 【0038】

【数6】

$$A = (I - \rho R)^{-1} C^* \quad \text{--- (6)}$$

ただし、 $I$ : 単位行列

【0039】なお、Anderson氏によれば、人間の連想記憶の特性を表すのに妥当な値として  $\rho^* = 1$ 、 $B = 0$ 、8であるとしている。これにより、キーワード連想ネッ

トワーク上の各ノード活性の最終的な値は、

【0040】

【数7】



$$A = 0.8 * (I - 0.8R)^{-1} C^*$$

--- (7)

ただし、R: リンクの強さ

C\*: 初期状態の活性値

により近似することができる。

【0041】また、請求項7のテキスト選定装置では、表示されたテキスト情報をユーザが読む際の視線の動きを追跡し、この追跡情報から表示されているテキスト情報に対するユーザの興味の度合を評価しているので、ユーザが持っている興味の度合をユーザが明示的に入力することなしにテキスト選定装置を学習させることが可能となる。

【0042】

【実施例】以下、図面を参照しながら実施例を説明する。図1には本発明の一実施例に係るテキスト選定装置のブロック構成図が示されている。

【0043】この装置は、テキストの選定を行う部分と、テキストの選定を行った後にそれをキーワード連想ネットワークに反映させるための学習を行う部分とに分けられる。

【0044】同図において、文書記憶部1は、新たに対象となるテキストを一時的に記憶しておくためのものである。この文書記憶部1に記憶されたテキストからキーワード抽出部2によって所定のキーワードが抽出される。

【0045】抽出されたキーワードはキーワード記憶部3に記憶される。このキーワード記憶部3に記憶されたキーワードについて、活性値計算部4によってキーワードの活性値が求められる。そして、求められた活性値はキーワード連想ネットワーク記憶部（以下「ネットワーク記憶部」という）5に記憶されたキーワード連想ネットワーク上に伝搬される。この伝搬結果に基づいてテキストの得点が計算される。閾値判定部6は、計算されたテキストの得点が所定の閾値以上か否かを判定し、閾値以上であれば、該テキストを表示すべく表示部7に送出する。また、表示部7は入力部8から入力されたデータやコマンド等を適宜表示する。キーワード連想ネットワーク更新部9は、入力部8により入力される後述する評価値に基づいてキーワード連想ネットワークを更新する。

【0046】次に、本実施例に係るテキスト選定装置の各部について詳述する。文書記憶部1に一時的に蓄えられた新たなテキストは、公知のキーワード抽出手段（たとえば、石川徹也：「文意解析処理に基づく主題索引作成支援システム」、情報処理学会論文誌Vol32, No. 2, 1991 に用いられているキーワード抽出手段）によってキーワードが抽出される。抽出されたキーワードはキーワード記憶部3に記憶される。

【0047】活性値計算部4は、まず、ネットワーク記

憶部5に記憶されているネットワーク上の各キーワードに対応するノード全てに初期値を与える。そして、対応するノードが存在する全てのキーワードについての処理が終了すると、活性値計算部4はネットワーク上に活性値を伝搬させて最終的な活性値を求める。さらに、各ノードの活性値と重みとを考慮した関数を用いて各ノードの得点を求める。本実施例においては、ノードの活性値と重みとの積をノードの得点としている。

【0048】ここで、本実施例において用いられるキーワード連想ネットワークについて説明する。図2はネットワークの一例を示す模式図である。

【0049】各キーワードに対応するノード21には、ユーザの得点の累計を表す「大きさ」と、ユーザにとっての重要度を表す「重み」とが設けられている。各ノード間はリンクが存在する場合と、存在しない場合とがある。リンクが存在する場合、ノードAからノードBへの「リンクの強さ」は、ノードAが直接リンクを持つ全てのノードの「大きさ」の合計でノードBの「大きさ」を割った値により決定する。すなわち、あるノードAから出る全てのリンクの強さの合計は常に1になる。

【0050】図3は、図2に示したネットワーク上のノード情報がネットワーク記憶部5に記憶されている状態の一例を示す図である。ノードの属性には「ノード名」、「大きさ」、「重み」、「活性値」および「得点」がある。本実施例では、活性値は初期状態において全て0である。

【0051】図4は、図2に示したネットワーク上のノード間のリンク情報がネットワーク記憶部5に記憶されている状態の一例を示す図である。行列のインデックス(i/j)は、図3におけるノード番号に対応する。ノード間にリンクが存在する場合は、行列の対応する要素を1で表し、リンクが存在しない場合は0で表す。ノードの自分自身へのリンクは存在せず、必ず0になる。また、ノードiからノードjへリンクが存在する場合は、同時にノードjからノードiへのリンクも存在するものとする。したがって、ノード間リンク情報を表す行列は必ず対称行列となる。なお、リンク情報は、必ずしも双方向である必要はなく、単方向のみを記述する三角行列であってもよい。

【0052】図5は、各ノード間の「リンクの強さ」に関する情報がネットワーク記憶部5に記憶されている状態の一例を示す図である。各ノード間のリンクの強さは、ノードiからノードjへのリンクの強さを表し、上記ノード情報とノード間のリンク情報とから求められる。たとえば、ノード「ユーザ」とノード「インターフ

10

20

30

40

50

エイス」とのリンクの強さ、すなわちノード番号  $i = 1$  からノード番号  $j = 2$  へのリンクの強さは、  
 リンクの強さ  $= 90 / (80 + 90)$   
 $= 0.53$   
 となる。

【0053】次に、活性値計算部4について説明する。ノード情報の「活性値」属性には、ネットワーク上の全てのノードの活性値が記憶されている。本実施例においては、初期状態は全て0である。活性値計算部4は、キーワード記憶部3よりキーワード群を読み出し、それぞれのキーワードに対応するノードがキーワード連想ネットワーク上に存在する場合には、該ノードの活性値を該ノードの「大きさ」と同じ値に設定する。全てのキーワードについて処理が終了すると、活性の伝搬を求める。あるノードの活性は減衰率  $\rho$  で定められる値により減衰した値がリンクを持つ隣のノードにリンクの強さに応じて分配される。この活性値計算部4の機能を図6に基づいて説明する。

【0054】図6は、活性値計算部4の機能を説明するための流れ図である。キーワード記憶部3には、図7に示すような抽出されたキーワードがキーワードリストとして記憶されているとする。活性値計算部4は、このキーワードリストを参照するためのインデックス  $i$  をキーワードリストの先頭である1に設定する（ステップ61）。

【0055】次に、キーワードリストに  $i$  番目のキーワードが存在するか否か、すなわちキーワードリストにまだ読み込まれていないキーワードが存在するか否かを検索し（ステップ62）、存在する場合は当該キーワードを読み込む（ステップ63）。読み込まれたキーワードが、ネットワーク記憶部5上に記憶されている場合（ステップ64）、対応するノードの活性値を該ノードの大きさと同じ値に設定する（ステップ65）。そして、インデックスを1つ進め（ステップ66）、これら一連の処理をキーワードリストの全てのキーワードについて行\*

$$0.15 \times 4 + 0.23 \times 4 + 0.17 \times 0 + 0.14 \times 3 + 0.16 \times 2 + 0.15 \times 1 = 2.41$$

となる。

【0060】一方、同じネットワークと図7(b)のキー※

$$0.13 \times 4 + 0.20 \times 4 + 0.30 \times 0 + 0.12 \times 3 + 0.14 \times 2 + 0.10 \times 1 = 2.06$$

となる。

【0061】次に、閾値判定部6について説明する。閾値判定部6は、テキストの得点が閾値以上であるときには表示部7に該テキストを表示する旨の制御命令を送る。したがって、表示部7に該テキストが表示される。テキストの得点が閾値以下であるときには表示は行われない。

【0062】このように、表示部7は閾値判定部6の制御命令にしたがって文書記憶部1に記憶されているテキストを表示する。したがって、上述の例で、閾値判定部6内に保持されている閾値が2.40と2.06との間、た

\*う。

【0056】キーワードリストの全てのキーワードについての処理を終了すると（ステップ62）、活性値計算部4は活性値の伝搬を求める（ステップ67）。ネットワーク上の各ノードの最終的な活性値は、既に述べたように、(7)式により近似される。(7)式の  $R$  は図5に示したノード間の「リンクの強さ」であり、 $C^*$  はテキストから抽出したキーワードに対応するノードがシステムにより活性化された結果、つまりネットワークに供給される活性値の量を表す。この活性値の量は、既にステップ65において「活性値」属性に代入されている。したがって、(7)式を計算して図3に示した「活性値」属性に順次書き込めば、ネットワークの活性値の伝搬を求めることと等価になる。なお、本実施例においては、(7)式の連立一次方程式を解くため、SOR (Successive Over-Relaxation) 法を用いているが、これに限る必要はない。

【0057】次に、キーワード連想ネットワークの学習に伴うノードの肥大による活性値のオーバーフロー（インフレーション）を避けるために活性値を正規化する。具体的には、全てのノードの活性の合計を求め、これを各ノードの活性の値で割り、これを図5の活性値の項に再び記憶する（ステップ68）。

【0058】次に、活性値計算部4は、活性値と重みとを考慮した関数により各ノードの得点を求める。ここでは、活性値と重みとの積により得点を求めるものとする。求められたノードの得点は、図5のノードの「得点」属性に書込まれる（ステップ69）。そして、テキストの得点は、全てのノードの得点を合計することにより求められる（ステップ70）。

【0059】たとえば、図2に示したネットワークと図7(a)に示したキーワードリストとが活性値計算部4に与えられた場合、活性値は図8(a)のように伝搬し、テキストの得点は、

※ワードリストとが活性値計算部4に与えられた場合、活性値は図8(b)のように伝搬し、テキストの得点は、

たとえば、2.20であった場合には、図7(a)のキーワードリストを生成したテキストはユーザに表示される。一方、図7(b)のキーワードリストを生成したテキストはユーザに表示されないことになる。

【0063】つまり、ネットワークに着目してみると、ユーザは「インタフェース」というキーワードに興味があることが分る。しかし、これは「ユーザインタフェース」の意味であり、ハードウェア間の「通信インタフェース」ではないので、ユーザの本来の意図に合致することになる。

【0064】上述の例ではどちらも「インタフェース」

というキーワードを含んでいたが、ノードの大きさが大きく、かつノードの重みが小さいノード（この例では0）を含んでいる場合よりも、ノードの大きさとノードの重みとが小さくても複数のノードで支持されている場合の方が、テキストの得点が高いという結果になる。

【0065】したがって、テキストから抽出されたキーワードリストに大きさと重みとが小さいノードが数多く含まれている場合であっても、テキストの得点として大きいものを得ることが可能となる。しかも、このような場合はユーザが本来ならば重みの大きいキーワードを単に知らない場合が多い。よって、このような方法によれば、新しいキーワードを発見することも可能である。逆に、重みの非常に大きいキーワードが一つだけキーワード群に含まれる場合は、通常のキーワード検索手段により達成することができるので、これと併用するようにしてもよい。

【0066】このように表示部7に表示されたテキストについて、ユーザはテキストを判定し、ユーザにとっての判定結果を評価値として入力部8より入力する。本実施例においては、0から4までの数字の5段階の評価値を用い、この数字が高い程ユーザにとって重要なテキストであることを表している。さらに、本実施例では、抽出されたキーワードであって、ネットワーク上に登録されていないものを表示部7に順次表示し、これらのキーワードの重みについても同様に5段階評価で入力部8よりユーザによって入力することを促すようにしている。

【0067】入力部8より入力された判定結果は、キーワード連想ネットワーク更新部9に送出される。キーワード連想ネットワーク更新部9では、キーワード記憶部3に記憶されているキーワードおよび入力された判定結果に基づいてネットワーク記憶部5に記憶されているネットワークを更新する。

【0068】次に、キーワード連想ネットワーク更新部9の機能について説明する。キーワード連想ネットワーク更新部9は、表示部7に表示されたテキストに対して入力部8から判定結果が与えられると、ネットワークの状態を更新する。

【0069】図9乃至図11は、キーワード連想ネットワーク更新部9の機能を説明するための流れ図である。図9に示す処理では、ネットワーク上に登録されていないキーワードに対する登録および今回新たにキーワード抽出部2が抽出したキーワードに対応するノードの大きさの変更を行っている。

【0070】すなわち、キーワード記憶部5に記憶されたキーワードリストを参照するためのインデックスiをキーワードリストの先頭である1に設定する（ステップ91）。次に、キーワードリストにi番目のキーワードが存在するか否かを確認し（ステップ92）、存在する場合は該当するキーワードを読み込む（ステップ93）。次に、この読み込まれたキーワードがネットワー

ク記憶部5のノード情報領域に記憶されていない場合（ステップ94）は、対応するノードをノード情報に追加し、「ノード名」属性に該キーワードを登録し、「大きさ」、「重み」、「活性値」およびノードの「得点」属性に0を代入し（ステップ95）、該キーワードを表示部7に表示する（ステップ96）。

【0071】ユーザは、該キーワードの重みを5段階で評価し、入力部8からその判定結果を入力する（ステップ97）。なお、この判定結果は、0から4の数値により表され、数値が大きいほどユーザにとって重要であるものとする。

【0072】次に、ノード情報の対応するノードの「重み」属性に該判定結果を代入する（ステップ98）。次に、先にユーザがテキストを評価したテキストの判定結果を対応するノードの「大きさ」属性に書き込む（ステップ99）。そして、キーワードのインデックスiを次に進める（ステップ100）。これら一連の処理をキーワードリストの全てのキーワードについて行って（ステップ92）、次の図10に示される処理に進む。

【0073】図10に示す処理では、今回新たに抽出した全てのキーワード間にリンクを持たせる。すなわち、キーワードリストのi番目のキーワードを読み込み（ステップ102）、読み込まれたキーワードのノード番号をキーワード記憶部5のノード情報領域から引き出し、変数aに代入する（ステップ103）。次に、リンクを張る相手先のキーワードのインデックスをjとし、このインデックスjがインデックスiの次を指し示すようにして（ステップ104）、キーワードリストにj番目のキーワードが存在するか確認し、存在する場合はキーワードリストのj番目のキーワードを読み込む（ステップ105）。読み込まれたキーワードのノード番号をノード情報領域から引き出し、変数bに代入する（ステップ106）。

【0074】次に、キーワード記憶部5のリンク情報領域において行列で表されるノード間のリンク情報の

(a, b) および (b, a) の要素をそれぞれ1に設定し（ステップ107）、キーワードリストのインデックスjを次に進める（ステップ108）。キーワードリストの全てのキーワードについてjに関しての処理が終了すると（ステップ105）、キーワードのインデックスiを次に進める（ステップ109）。このようにして、一連の処理をキーワードリストの全てのキーワードについて行って（ステップ102）、次の図11に示される処理に進む。

【0075】図11に示す処理では、ノード間のリンクおよびノードの大きさの変更に伴い、図5に示したリンク間の強さの変更を行う。すなわち、ノード情報を参照するためのインデックスiをノード情報の先頭である1に設定する（ステップ111）。次に、ノード情報にi番目のノードが存在するか確認する（ステップ11

2)。リンクの強さは、その方向によりそれぞれ異なるため、ここにいうノード  $i$  はリンクの出発地点を表す。次に、ノード情報を参照するためのインデックス  $j$  をノード情報の先頭である 1 に設定し、ノードの大きさの合計を求める変数「合計」を初期設定のため 0 にする（ステップ 113）。

【0076】次に、ノード情報に  $j$  番目のノードが存在するか確認する（ステップ 114）。ここで、ノード  $j$  はリンクの目的地を表す。キーワード記憶部 5 のリンク情報領域においてリンクの有無を表す行列の各要素を「リンク  $[i, j]$ 」で表すものとし、リンク  $[i, j] \times$  大きさ  $[j]$  を変数「合計」に書き加えることにより、リンクが存在する場合は接続されているノードの大きさを書き加えることができる（ステップ 115）。次に、ノード情報に対するインデックス  $j$  を次に進める（ステップ 116）。インデックス  $j$  に関して全てのノードを処理し終えたら（ステップ 114）、インデックス  $j$  を再びノード情報の先頭である 1 に戻す（ステップ 117）。この時点で、ノード  $i$  に接続されている全てのノードの大きさの合計が求まることになる。

【0077】次に、ノード情報に  $i$  番目のノードが存在するか確認し（ステップ 118）、ノードが存在する場合は、図 5 に示したノード間の強さの行列の要素「強さ  $[i, j]$ 」に大きさ  $[j]$  の値の合計で割った値を代入する。ただし、リンク  $[i, j]$  を掛けることにより、リンクが存在しない場合には強さが 0 になるようにする（ステップ 119）。

【0078】次に、ノード情報のインデックス  $j$  を次に進める（ステップ 120）。ノード情報の全てのノードを  $j$  に関して処理し終わると（ステップ 118）、ノード情報のインデックス  $i$  を次に進める（ステップ 121）。ノード情報の全てのノードを  $i$  に関して処理し終わると（ステップ 112）、処理を終了する。

【0079】図 2 に示したキーワード連想ネットワークに、図 7 (a) のキーワード群を生成するテキストの得点が 4 で、重みが 4 の新規キーワード「マン・マシン」を加えた更新結果を図 12 に示す。新たに加わったリンクを実線で示し、既に存在しているリンクを点線で示している。

【0080】ところで、上記例では活性値計算部 4 で、ACT\* の方法を使って活性値を伝搬させ、キーワード連想ネットワーク上の最終的な活性値を求めるようにしている。ACT\* の方法をそのまま適用すると、キーワード連想ネットワーク上に存在するノードの数と同じ次元の連立一次方程式を解く必要がある。連立一次方程式を解くには、その方程式の次元の三乗に比例して計算量が増加するため、大規模なキーワード連想ネットワークに用いるのは困難が伴う。

【0081】そこで、ここでは大規模なキーワード連想ネットワークにおいても、計算爆発を起こさずに活性伝

搬の定常状態を求めることができる二通りの活性伝搬手法を説明する。

#### 【0082】(1) 単純な距離を用いる方法

この方法では、テキストから抽出された各キーワードについて、各キーワードを中心とし、キーワード連想ネットワークのリンクをたどって、決められたステップ数以内で到達可能なノードとリンクのみを取り出した、部分ネットワークと ACT\* の式とを用いて活性値を求める。そして、求まった部分ネットワークの各ノードの活性値をキーワード連想ネットワークの対応するノードに加え、これを全てのキーワードに関して行なうことにより、キーワード連想ネットワークの全ノードの活性値を求める。

【0083】以下、図 13 に示す処理の流れを参照しながら説明する。まず、図 1 のキーワード記憶装置 3 に記憶されている図 7 に示すようなキーワードリストを参照するためのインデックス  $i$  をキーワードリストの先頭である 1 に設定する（ステップ 131）。次にキーワードリストに  $i$  番目のキーワードが存在するか確認し（ステップ 132）、存在する場合にはキーワードリストの  $i$  番目のキーワードを読み込む（ステップ 133）。読み込まれたキーワードが、図 3 のようなキーワード連想ネットワークに記憶されている場合（ステップ 134）、変数“起点”を読み込まれたキーワードに設定し、変数“距離”を予め定められた部分ネットワークの大きさ（この例では 5）に設定する（ステップ 135）。

【0084】図 17 のような部分ネットワークの初期化を行なう（ステップ 136）。ただし、初期化された段階では部分ネットワークには、どのノードも登録されていない。部分ネットワーク生成部に引数（起点、距離）を渡し、起点を中心としてリンクを辿って、5 つ隣のノードまでも含む部分ネットワークを作成する（ステップ 137）。

【0085】得られた部分ネットワーク内の起点の活性値をキーワード連想ネットワーク（全体ネットワーク）での起点のノードの大きさに設定する（ステップ 138）。部分ネットワークに含まれるノード間のリンクの強さ情報を図 5 のようなリンク情報から取り出し（ステップ 139）、活性の伝搬を求める（ステップ 140）。部分ネットワークの各ノードの活性値を対応する全体ネットワークの各ノードの活性値に加える（ステップ 141）。インデックス  $i$  を 1 つ進め（ステップ 142）、次のキーワードの処理へ進む。全てのキーワードの処理が終了したら（ステップ 132）、処理を終了する。

【0086】次に、図 14 を用いて部分ネットワーク生成部の働きを説明する。部分ネットワーク生成部は引数（距離）が正である間、再帰的に部分ネットワークを拡張して行く。まず、引数として受け取った起点の全体ネットワークでの番号を部分ネットワークに登録し、その

活性値を0に設定する(ステップ141)。次に、引数として受け取った距離を1減じ、その値を変数“次の距離”に格納する(ステップ142)。次の距離が0であった場合にはシステムは終了し、呼び出し元に戻る(ステップ143)。次の距離が0でなかった場合にはインデックスとなる変数*i*を0に初期化する(ステップ144)。図4のようなリンク情報を参照し、起点ノードに隣接しているノード、隣接ノード[i]を探し、隣接ノードを全て探し終えたら終了する(ステップ145)。

【0087】隣接ノード[i]が部分ネットワークに登録されていなければ(ステップ146)、隣接ノード[i]とを引数として、再帰的に部分ネットワーク生成部を呼ぶ(ステップ48)。隣接ノード[i]が部分ネットワークに登録されていれば(ステップ146)、隣接ノード[i]の距離と次の距離とを比較し、次の距離の方が大きければ(ステップ147)、隣接ノード[i]と次の距離を引数として、再帰的に部分ネットワーク生成部を呼ぶ(ステップ148)。隣接ノードを探すインデックスを一つ進め、次の隣接ノードを探す(ステップ149)、隣接ノード[i]の距離が次の距離以上の場合には、隣接ノードを探すインデックスを一つ進め、次の隣接ノードを探す(ステップ149)。

【0088】このようにして、リンクをたどって、始めの距離で定められた回数で到達可能な範囲に存在するノードを全て取り出した、部分ネットワークを生成する。次に、図15を用いて部分ネットワークに対応したノード間のリンクの強さの情報を作成する方法について説明する。

【0089】まず、インデックス*i*を1に初期化する(ステップ151)。インデックス*i*(部分ネットワークに含まれるノードの数)がN以下の場合にはノード[i]に関して処理を行ない、Nに到達した、すなわち部分ネットワークに含まれる全てのノードに関して処理を終えた場合には終了する(ステップ152)。

【0090】インデックス*j*を1に初期化する(ステップ153)。インデックス*j*(部分ネットワークに含まれるノードの数)がN以下の場合にはノード[j]に関して処理を行ない(ステップ154)、Nに到達した、すなわち部分ネットワークに含まれる全てのノードに関して処理を終えた場合にはインデックス*i*を1つ進め、再び全ての*j*に関して処理を行なう(ステップ155)。

【0091】変数*a*にノード[i]の全体のネットワークでの番号を格納し、変数*b*にノード[i]の全体のネットワークでの番号を格納する(ステップ156)。ノード[i]からノード[j]への強さ:部分ネットワークの強さ[i, j]は強さ[a, b]であるため、この値を格納する(ステップ157)。インデックス*j*を1つ進め(ステップ158)、次のノードに関して処理を続ける。

【0092】このようにして全てのノードに関して処理

を行なうと、必要なリンク情報の取り出しが完了する。

(2) リンクの強さを考慮した距離を用いる方法

この方法では、テキストから抽出された各キーワードについて、各キーワードを中心とし、キーワード連想ネットワークのリンクをマーカパッシング手法を用いて、リンクの強さに応じてマーカのzorchの値を減少させ、最終的にzorchが到達した範囲までのノードとリンクのみを取り出した、部分ネットワークとACT\*の式とを用いて活性値を求める。求まった部分ネットワークの各ノードの活性値をキーワード連想ネットワークの対応するノードに加える。これを、全てのキーワードに関して行なうことにより、キーワード連想ネットワークの全ノードの活性値を求める。

【0093】すなわち、先の(1)の例では隣接するノードを一つ進む毎に距離を一つ減らしたが、活性の伝搬はリンクの強さに依存するため、この方法ではリンクの強さを考慮した関数を用いて距離を算出する。これは先に説明した(1)の例の部分ネットワーク作成部を置き換えることにより実現可能である。

【0094】部分ネットワークの作成部の一例を図16を用いて説明する。先の(1)の例での部分ネットワーク作成部と同様に再帰的に処理を行なう。引数として与えられた起点を図17のような部分ネットワークに登録する(ステップ161)。インデックス*i*を0に初期化する(ステップ162)。処理していない起点に隣接するノードがまだ存在する場合には処理を行なう(ステップ163)。

【0095】隣接するノードと起点の距離を距離関数を用いて求める。この例では一例として、起点ノードから隣接ノードへのリンクの強さの逆数を距離として用いる。すなわち、 $1/\text{強さ}[\text{起点}, \text{隣接ノード}]$ を起点から隣接ノードへの距離とする。起点から隣接ノードへの距離の値を変数「距離」から引いた値を変数として次の距離に格納し(ステップ164)、次の距離が負でないならば、処理を行なう(ステップ165)。隣接ノードが既に部分ネットワークとして登録されている場合には(ステップ166)、次の距離が登録されたときの距離よりも大きい場合にのみ(ステップ167)、部分ネットワーク作成部を再帰的に呼び出すことにより更新を行なう(ステップ168)。隣接ノードが登録されていない場合には(ステップ166)、部分ネットワーク作成部を再帰的に呼び出すことにより登録を行なう(ステップ168)。インデックス*i*を1つ進めて(ステップ169)、次の隣接ノードの処理を行なう。

【0096】なお、上記手法に加え、さらに活性値情報を全体ネットワークの一部として持たせるのではなく、図18のような活性値情報テーブルを独立に持たせ、活性値に変化のあったノードの活性値だけを記憶させることにより、文書の得点を計算する際に活性が初期状態(0)のままのノードは考慮する必要がなくなり、文書

の得点計算を高速化することが可能となる。

【0097】すなわち、部分ネットワークで求めた各ノードの活性値情報を図18のような活性ノードテーブルに登録することにより、活性値が正であるノードのみの一覧を作成することができる。

【0098】すなわち、図13のステップ140において、部分ネットワークを全体ネットワークに加える際に、ノード番号と活性値を図18のような活性ノードテーブルに登録する。既にノードが登録されている場合には、部分ネットワークで新たに求めた活性値を足す。文書の得点を求める際には活性ノードテーブルを参照し、先頭ノードから順に、ノード番号を求め、ノード番号を用いて対応するノードの重みを全体ネットワークから得る。次に、活性値を得て、この2つの値から、ノードの得点を得る。上記の処理を活性ノードテーブルに登録されている全てのノードに関して行なうことにより、文書の得点を得ることができる。

【0099】また、図1に示した実施例では、表示部7に表示されたテキストをユーザが読みながら選定された結果をユーザ自身が5段階で評価し、この評価値をユーザが直接入力部8を介して入力する方法を採用している。しかし、このような評価入力手法では、迅速性および的確性を満たすことが困難となる。したがって、選定結果を簡単かつ的確に評価する手法の出現が望まれる。

【0100】提示されたテキストを評価する手法には、特開平4-192751号公報に示されている個人向け電子新聞システムに見られるような、探索や操作の履歴を用いて検索などの操作が行われた情報に興味があると判定する方法や、特開平4-77866号公報に示されている情報提供システムの情報提供方法に見られるような、情報の提示量と提示時間を元に情報単位あたりの提示時間から利用者の興味を算出する方法などが知られている。

【0101】しかし、検索や操作の履歴を用いる方法では間接的にしかユーザの興味の度合いが求められず、また情報に対して興味はあったが、それ以上の検索や操作をせずに満足してしまったような情報については興味がなかった情報と区別することが不可能である。また、情報の提示量と提示時間を元に、情報単位あたりの提示時間からユーザの興味を算出する方法では、情報が提示された状態で放置されてしまった場合や、考え事をしてしまったような場合には、たとえ興味がなくとも重大な興味があったと判定されてしまう虞がある。

【0102】そこで、ここでは、図1に示すテキスト選定装置によって選定され、表示されているテキストを見るユーザ、つまり利用者の視点を追跡し、その移動特性を解析することにより、提示されているテキストに対して利用者が持っている興味の度合いを利用者が明示的に入力することなく、迅速かつ的確に評価する評価装置について説明する。

【0103】すなわち、この評価装置は、図1に示されているテキスト選定装置により重要と判定され、表示されているテキストを利用者が読む際の視線の動きを追跡する視線追跡装置と、この視線追跡装置により得られた座標情報をテキストにおける位置情報に変換する座標変換装置と、変換された位置情報を用いて利用者の興味の度合いを評価する評価値算出装置と、評価の際に用いる係数と過去の予測値と評価値とを記憶する評価値記憶装置と、記憶された予測値と評価値とを用いて係数の学習を行なう学習装置とを備えている。

【0104】この評価装置では、表示されたテキストを利用者が見る際に、Applied Science LaboratoriesのModel 3250Rのような視線（視点）追跡装置を用いて表示装置上の視点の位置を常に記録する。人間は興味によって表示装置上の視点の運び方が異なる。すなわち、テキスト情報を表示している場合、興味のある情報ならば、視点の動きに乱れが少なく、文字を先頭から順次追う。逆に興味のない情報であれば、飛ばしながら読んだり、途中で止めたり、視点が表示装置外に移動する等の乱れが発生する。この評価装置では、これらの乱れを検出し、乱れの度合いにより提示テキストに対する興味の度合いを評価する。

【0105】測定された興味の度合いは、図1に示したテキスト選定装置の入力部8に渡され、ユーザ情報として管理、利用される。なお、この評価装置で測定された興味の度合いは、図1に示したテキスト選定装置に限らず、他のタイプのテキスト選定装置、たとえば特開平2-125363号公報の文書検索装置に見られるようなキーワードコネクションによるものにも利用できる。この文書検索装置では、キーワード間のつながりの強さをネットワークで表現し、検索条件からのつながりの強さを基に文書検索を行ない、検索結果を利用者が判定／入力し、判定結果を用いてキーワード間のつながりを学習させ、次の検索からより利用者の意図に近い挙動を取るように変化する。

【0106】肉体的衰えや学習、訓練により視点の運びは徐々に変化する。したがって、視点の動きに対応する評価値を徐々に変化、学習させる仕組みが必要である。図1に示されるテキスト選定装置は、情報を提示する以外にもファイリング、情報検索、仮名漢字変換の辞書などにも利用され、異なるフィードバック情報により学習することも可能であるため、このような装置に評価予測機能を持たせた場合、予測と実際の評価値が連続して大きく異なる場合には、評価装置が現実とずれてしまったと判定し、評価装置の評価に用いる係数を現実に合わせて学習させる必要がある。

【0107】図19には評価装置の一例が示されている。ここで、評価装置より得られた評価値は図1の入力部8を介してテキスト選定装置に渡される。そして、この図19では図1に示されるテキスト選定装置を番号3

2で示してある。

【0108】文書記憶部31から入力されたテキスト情報は、テキスト選定装置32により利用者にとってどの程度重要であるか判定され、重要であると判定されたものに関してのみ点数付けされ、点数と共に情報記憶装置33に記憶される。

【0109】制御装置34はテキスト選定装置32により付加された点数を基に、情報記録装置33より利用者にとって有用と思われるテキスト情報を取り出し、入力装置35からの要求に合わせて表示装置36にテキスト情報を表示する。

【0110】視線追跡装置37は表示装置36に表示されているテキスト情報を読む利用者の視線の動きを追跡し、表示装置36上の利用者の視点の位置を座標情報として座標変換装置38に渡す。

【0111】座標変換装置38は、情報記憶装置33と制御装置34より表示装置36に表示されているテキスト情報を獲得し、視線追跡装置37より得た表示装置36上の利用者の視点の位置をテキスト情報の座標に変換する。

【0112】入力装置35より次文書表示命令もしくは終了命令が入力されると、変換された座標情報はテキスト情報と共に評価値算出装置39に渡される。評価値算出装置39は、評価値記憶装置40から評価に用いる係数を読み込み、該情報から利用者の興味の度合を判定し、算出された評価値をテキスト選定装置32にフィードバック情報として渡す。テキスト選定装置32は、このフィードバック情報を基に学習を行なう。

【0113】学習装置41は、情報記憶装置33からテキスト選定装置32が付加した点数と評価値算出装置39が行なった評価の結果とを受け取り、評価値記憶装置40に保存し、両者が大きく異なる場合には評価値算出装置39の学習を行ない、評価値記憶装置40に学習後の係数を保存する。

【0114】次に、図20を参照しながら制御装置34の働きを説明する。情報記憶装置33より文書情報とテキスト選定装置32が付加した得点とを読み込み、座標変換装置38に読み込んだ文書の番号を通知する。(ステップ211)。変数である表示開始行を1に初期化する(ステップ212)。表示開始行からN行分の文書情報を表示装置36に表示する(ステップ213)。ただし、Nとは表示装置36に一度に表示することが可能な行数であるとする。

【0115】表示開始行を座標変換装置38に通知し(ステップ214)、入力装置35からの命令の入力を待つ(ステップ215)。入力があった場合、入力が次ページ表示命令ならば(ステップ216)、表示開始行をNだけ進めてステップ213に戻る(ステップ217)。入力が次文書表示命令ならば(ステップ218)、評価開始を座標変換装置38に通知し(ステップ

219)、ステップ211に戻る。入力が終了命令ならば(ステップ220)、評価開始を座標変換38に通知し(ステップ221)、終了する。

【0116】次に、図21を参照しながら座標変換装置38の働きを説明する。制御装置34より現在表示している文書の文書番号を受け取る(ステップ231)。受け取った文書番号を基に情報記憶装置33より該当文書を読み込む(ステップ232)。制御装置34より表示開始行を受け取る(ステップ233)。

10 【0117】制御装置34から命令がない場合には(ステップ234)、視線追跡装置37から表示装置36上の利用者の視点の座標情報を受け取り(ステップ235)、表示開始行情報を基に文書のどの位置を読んでいるかを判定し、図24のような形式の文書上の座標に変換し(ステップ236)、ステップ234に戻る。

20 【0118】制御装置34から命令がある場合(ステップ234)、その命令が表示開始行の変更命令である場合には(ステップ237)、表示開始行を変更し(ステップ238)、ステップ234に戻る。その命令が表示開始行の変更命令でない場合で(ステップ237)、その命令が評価開始命令でない場合には(ステップ239)、ステップ234に戻る。その命令が評価開始命令である場合には(ステップ239)、評価値算出装置38を起動し(ステップ240)終了する。

【0119】次に、実例を用いて座標変換例を説明する。表示装置36に図22に示すような文書の一部が表示されているとする。ここでは6行目から表示されているので表示開始行は6である。また、表示可能な行数

30 (N)は12である。視線追跡装置37から表示装置36の左上を(0, 0)とした座標が入力として一定間隔で座標変換装置38に渡される。ここでは表示装置36の左上を(0, 0)、右下を(100, 100)とする座標が渡されるものとし、利用者が表示装置36以外の場所を見た場合は無視する。

【0120】今、視線追跡装置37から図23に示すような入力があったとする。この場合、表示装置36上の視点の位置は図25中に「×」で示す位置であり、視点は図26に実線で示すように移動したと考えられる。

40 【0121】座標変換装置38では、このような座標情報を文書の位置情報に変換する。上の例では座標(4, 2)は[6, 1]に変換する。ただし[行、文字数]で[6, 1]は6行目の1文字目を意味する。同様にして、図24のように全ての座標を変換して評価値算出装置39に渡す。

【0122】評価値算出装置39では変換後の座標情報を用いて利用者がどの程度表示文書に対して興味を持ったかを評価する。この評価の方法には様々な方法が考えられる。たとえば、利用者が余りにも早く目を通すと、それは内容を理解していないと判定し、文字を処理する速度が一定速度以内である確率を算出し、これを評価値

とする方法や、文書の領域をどれだけくまなく見たかを基に、その割合いを評価値とする方法、文書を読んでいる間に表示装置36の領域外に視点が移動した回数を用いて、回数が多くなれば興味が薄いと評価する方法、視点の動きが文字の並びと逆行する回数を用いて、回数が多くなれば興味が薄いと評価する方法や、さらにこれらの組合せにより数多くの方法が考えられる。また、これらの評価にニューラルネットワークを用いることも可能である。

【0123】ここでは、一例として、文書の領域をどれだけくまなく見たかを基に、その割合いを評価値としている。図27には上記観点で評価値を算出する処理の流れが示されている。

【0124】まず、現在対象としている文書を情報記憶装置33より読み込み（ステップ331）、全文字数を数える（ステップ332）。利用者が目を通した文字をカウントするための変数である処理済み文字数を0に初期化する（ステップ333）し、一つ前の座標を記憶しておくための変数である前の座標を〔0, 0〕に初期化する（ステップ334）。

【0125】入力座標がまだある場合には（ステップ335）、次の座標を一つ読み込む（ステップ336）。読み込んだ座標が前の座標と同じ行にある場合には（ステップ337）、この二つの座標間の文字を読んだとみなし、この間にある文字数を処理済み文字数に加える（ステップ338）。

【0126】読み込んだ座標を変数である前の座標に代入し（ステップ339）、ステップ335に戻る。ステップ335で次の座標が存在しない場合には、評価値記憶装置37より係数を読み込み（ステップ340）、評価値を求める（ステップ341）。求めた評価値を学習装置38とテキスト選定装置32に出力して（ステップ342）、終了する。

【0127】たとえば全文字数が100の文書で、図28中に実線で示すように視線が動いた場合には、処理済みの文字数は90となり、評価値は $(90/100) \times$ 係数となる。たとえば係数が5だとすると、評価値は4.5となる。一方、同じ文書を図29中に実線で示すように視線が動いた場合には、処理済みの文字数は50となり、評価値は $(50/100) \times$ 係数となる。係数が5であれば、評価値は2.5となり、利用者は図28の場合に比べて興味が無かったと評価される。

【0128】学習装置41は利用者モデル32からの予測値と評価値算出装置39からの評価値を決められた回数分（H回）保持し、これらの値の差が連続して一定の閾値を越える場合には、係数を学習させ、差が各回において半分になるように係数を設定する。すなわち、予測値と評価値の差が全て半分となるような係数を新たに選ぶ。

【0129】これは、新しい係数＝〔現在の係数（予測

値－評価値）／2評価値〕＋現在の係数により求まる。これをH回全ての予測値と評価値の組に対して求め、その平均を新しい係数とする。たとえば、係数が5でHが3、閾値が2だとして、

- ・予測値5，評価値2
- ・予測値4，評価値1
- ・予測値4，評価値2

という組合せが連続して記録された場合、上記の式に当てはめ、平均を取ると新しい係数は18.75となり、同じ入力によってより大きな評価が得られるように修正される。なお、ここで説明した評価装置はテキスト情報に限らず、図形やビットマップによる絵などの評価にも利用することが可能である。

#### 【0130】

【発明の効果】本発明によれば、膨大な量のテキストの中からユーザにとって有益なもののみを選定する際に、従来のテキスト選定装置では困難とされた、さほど重要でないキーワードであるが、それらが数多く出現した場合には、それをユーザにとって必要とするテキストであるとして抽出することができる。すなわち、個々のユーザによって相違するキーワード相互間の類似性を柔軟に表現することができ、所望のテキストのみを選定することができる。

#### 【図面の簡単な説明】

【図1】本発明の一実施例に係るテキスト選定装置の構成図

【図2】同装置のキーワード連想ネットワーク記憶部に記憶されているキーワード連想ネットワークの一例を示す模式図

【図3】同装置のキーワード連想ネットワーク記憶部に記憶されたノード情報の一例を示す図

【図4】同装置のキーワード連想ネットワーク記憶部に記憶されたノード間のリンク情報の一例を示す図

【図5】同装置のキーワード連想ネットワーク記憶部に記憶されたノード間のリンクの強さに関する情報の一例を示す図

【図6】同装置の活性値計算部の機能を説明するための流れ図。

【図7】同装置のキーワード記憶部に記憶されたキーワードの一例を示す図

【図8】活性化されたキーワード連想ネットワークの一例を示す模式図

【図9】同装置のネットワーク更新部の機能を説明するための流れ図

【図10】同装置のネットワーク更新部の機能を説明するための流れ図

【図11】同装置のネットワーク更新部の機能を説明するための流れ図

【図12】同装置のネットワーク更新部により更新されたキーワード連想ネットワークの一例を示す模式図



【図 13】部分ネットワークに限定して活性値を算出する位置例における処理の流れ図

【図 14】同部分ネットワークを作成するときの処理の流れ図

【図 15】同部分ネットワーク内のノード同士を結ぶリンクの強さ情報を作成するときの処理の流れ図

【図 16】別の部分ネットワークを作成するときの処理の流れ図

【図 17】部分ネットワークの登録例の一例を示す図

【図 18】部分ネットワークに対応した活性ノードテーブルの一例を示す図

【図 19】評価装置の一例を示す構成図

【図 20】同評価装置における制御装置の動作を示す流れ図

【図 21】同評価装置における座標変換装置の動作を示す流れ図

【図 22】表示装置への表示例を示す図

【図 23】座標変換装置への入力座標の例を示す図

【図 24】座標変換装置での変化出力の例を示す図

【図 25】表示装置上での視点位置を説明するための図

【図 26】表示装置上での視点位置の動きを説明するための図

【図 27】評価値を得るための処理の流れ図

【図 28】利用者が文書に興味がある場合の視線の動き \*

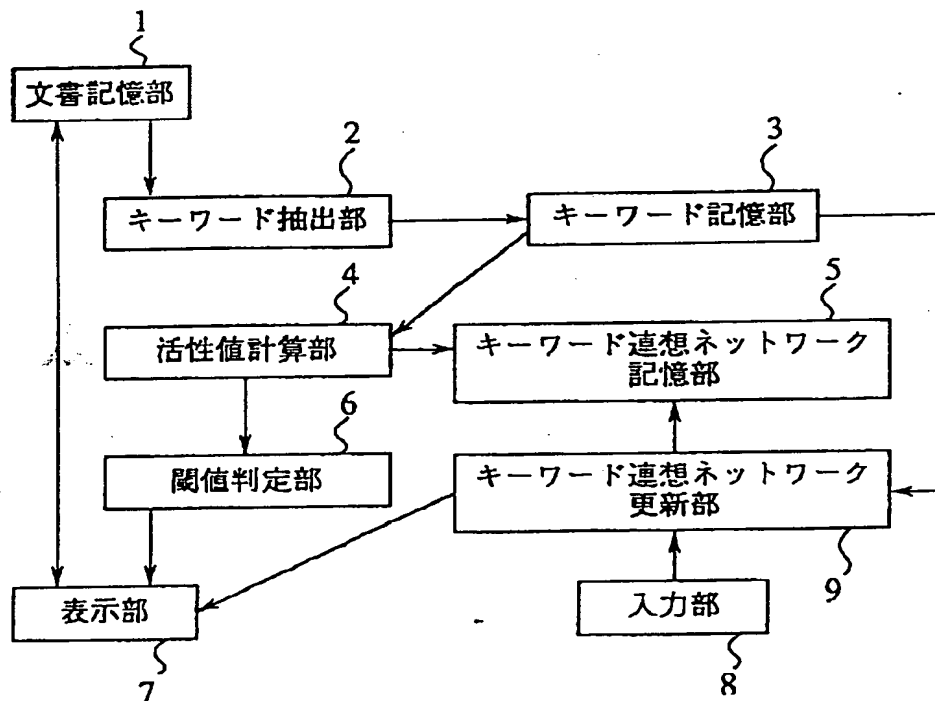
\* を説明するための図

【図 29】利用者が文書に興味がない場合の視線の動きを説明するための図

【符号の説明】

- |                    |        |
|--------------------|--------|
| 1…文書記憶部            | 2…キーワー |
| ド抽出部               |        |
| 3…キーワード記憶部         | 4…活性値計 |
| 算部                 |        |
| 5…キーワード連想ネットワーク記憶部 | 6…閾値判定 |
| 部                  |        |
| 7…表示部              | 8…入力部  |
| 9…キーワード連想ネットワーク更新部 | 21…ノード |
| 31…情報入力装置          | 32…テキス |
| ト選定装置              |        |
| 33…情報記憶装置          | 34…制御装 |
| 置                  |        |
| 35…入力装置            | 36…表示装 |
| 置                  |        |
| 37…視線追跡装置          | 38…座標変 |
| 換装置                |        |
| 39…評価値算出装置         | 40…評価値 |
| 記憶装置               |        |
| 41…学習装置            |        |

【図 1】



【図 7】

(1)	マン・マシン
(2)	インタフェース
(3)	マウス
(4)	アイコン

(a)

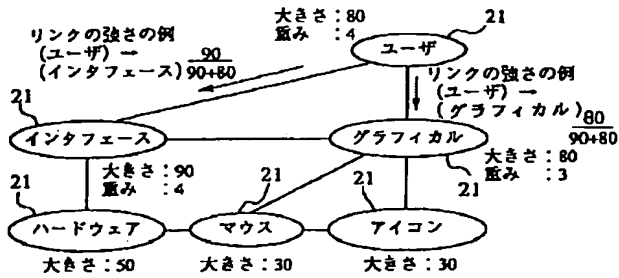
(1)	ハードウェア
(2)	インタフェース
(3)	通信
(4)	プロトコル

(b)

【図 23】 【図 24】

$\begin{pmatrix} 4. & 2 \\ 5. & 57 \\ 3. & 94 \\ 11. & 18 \end{pmatrix}$ 
 $\begin{pmatrix} 6. & 1 \\ 6. & 12 \\ 6. & 19 \\ 7. & 4 \end{pmatrix}$

【図 2】



【図 3】

番号	ノード名	大きさ	重み	活性値	得点
(1)	ユーザ	80	4	0	0
(2)	インタフェース	90	4	0	0
(3)	ハードウェア	50	0	0	0
(4)	グラフィカル	80	3	0	0
(5)	マウス	30	2	0	0
(6)	アイコン	30	1	0	0

【図 18】

【図 4】

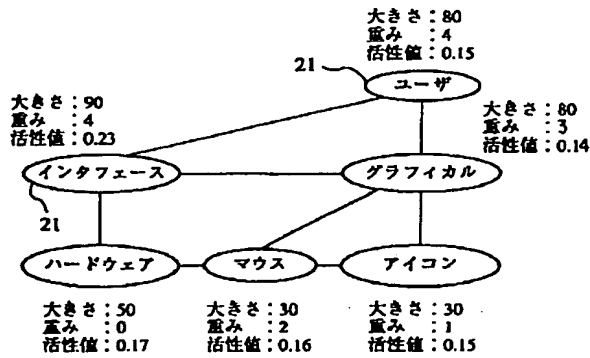
i/j	(1)	(2)	(3)	(4)	(5)	(6)
(1)	0	1	0	1	0	0
(2)	1	0	1	1	0	0
(3)	0	1	0	0	1	0
(4)	1	1	0	0	1	1
(5)	0	0	1	1	0	1
(6)	0	0	0	1	1	0

【図 5】

i/j	(1)	(2)	(3)	(4)	(5)	(6)
(1)	0.00	0.53	0.00	0.47	0.00	0.00
(2)	0.38	0.00	0.24	0.38	0.00	0.00
(3)	0.00	0.75	0.00	0.00	0.25	0.00
(4)	0.35	0.39	0.00	0.00	0.13	0.13
(5)	0.00	0.00	0.31	0.30	0.00	0.19
(6)	0.00	0.00	0.00	0.73	0.27	0.00

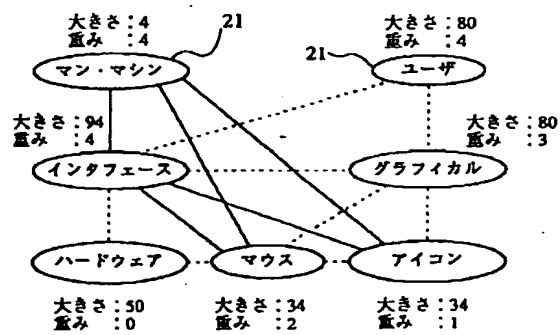
番号	ノード番号	活性値
(1)	1	0.15
(2)	4	0.14
(3)	3	0.23
.	.	.
.	.	.
(N)	5	0.14

【図 8】



(a)

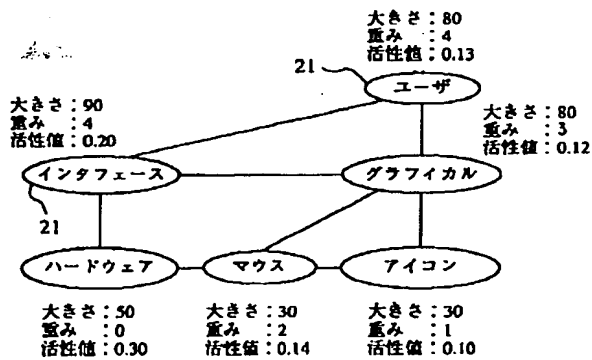
【図 12】



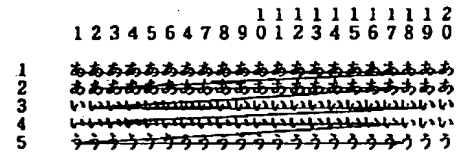
【図 17】

番号	ノード番号	距離
(1)	1	5
(2)	4	4
(3)	3	4
.	.	.
.	.	.
(N)	5	1

(b)

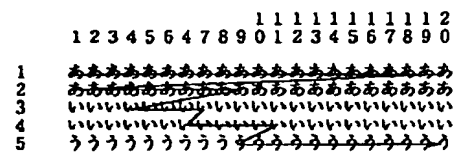


【图 28】



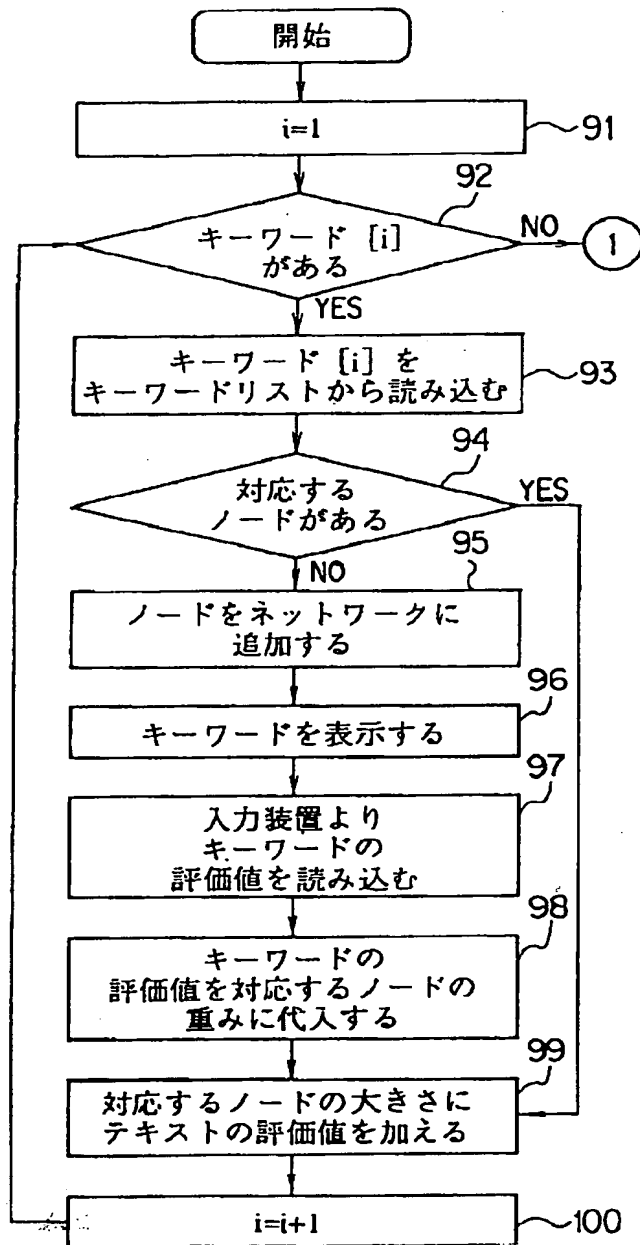
利用者が文書に興味がある場合の視組の選びの例

【图 29】

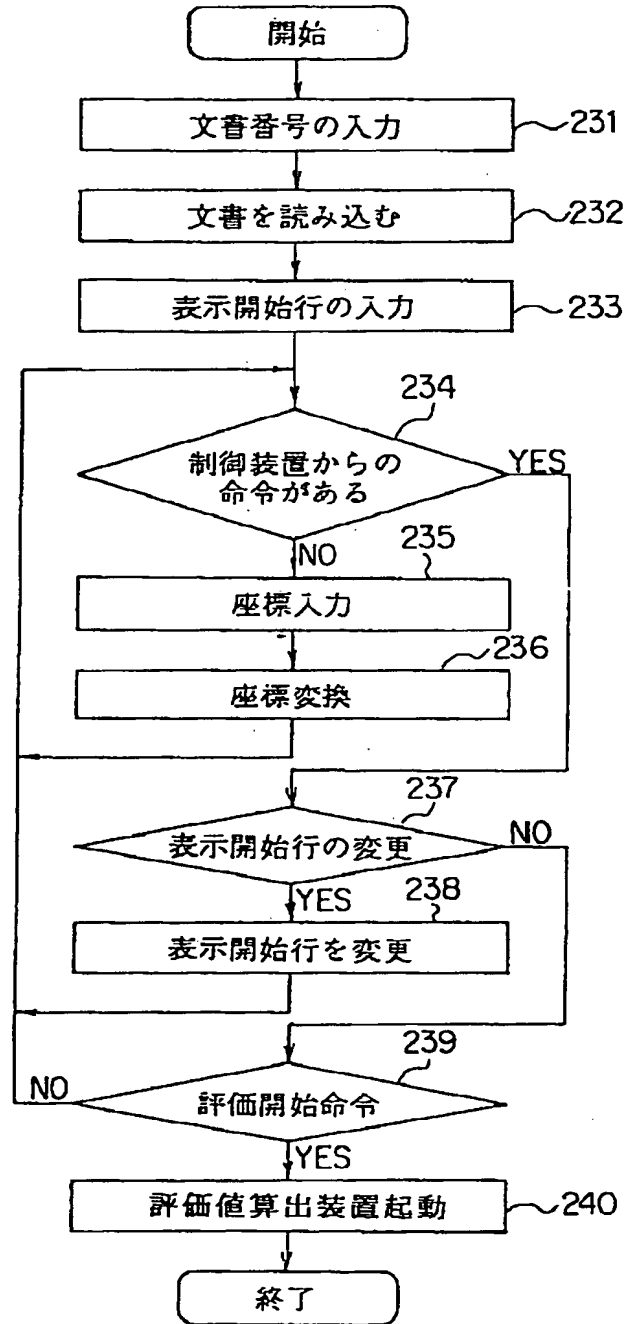


### 利用者が文書に興味がない場合の視線の運びの例

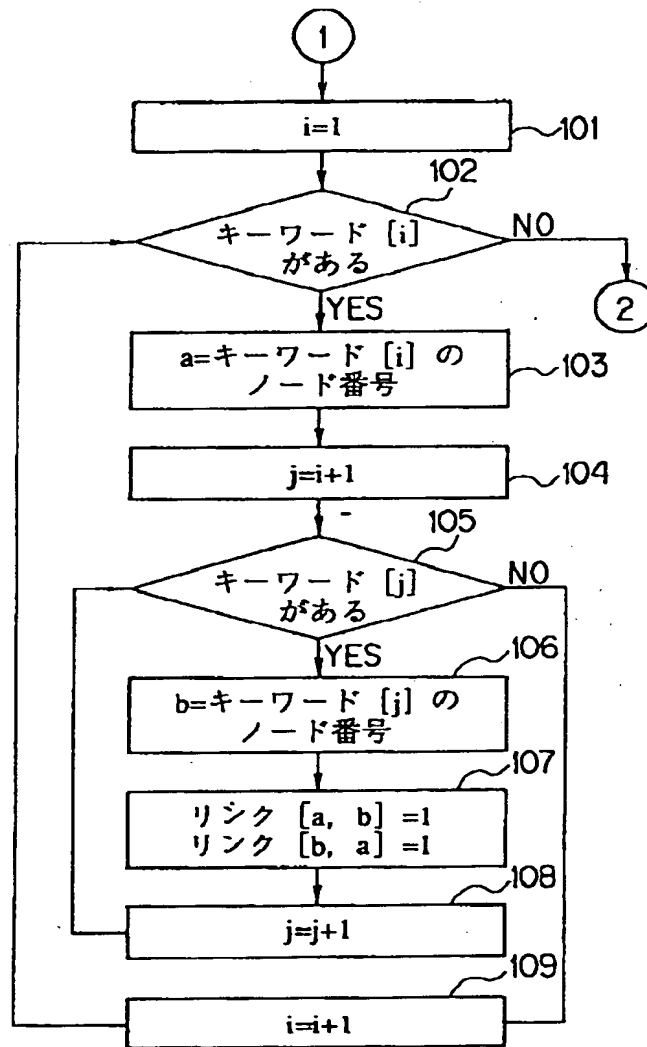
【図 9】



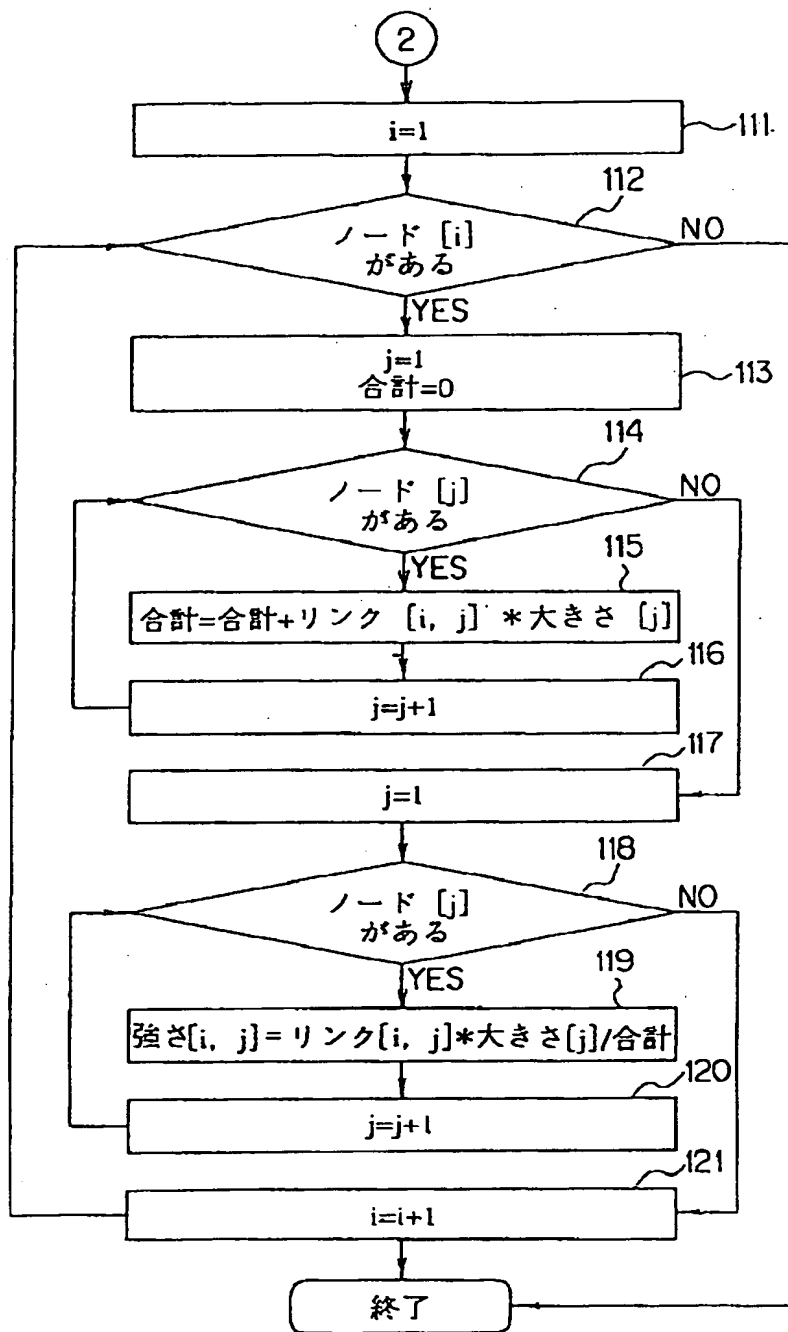
【図 21】



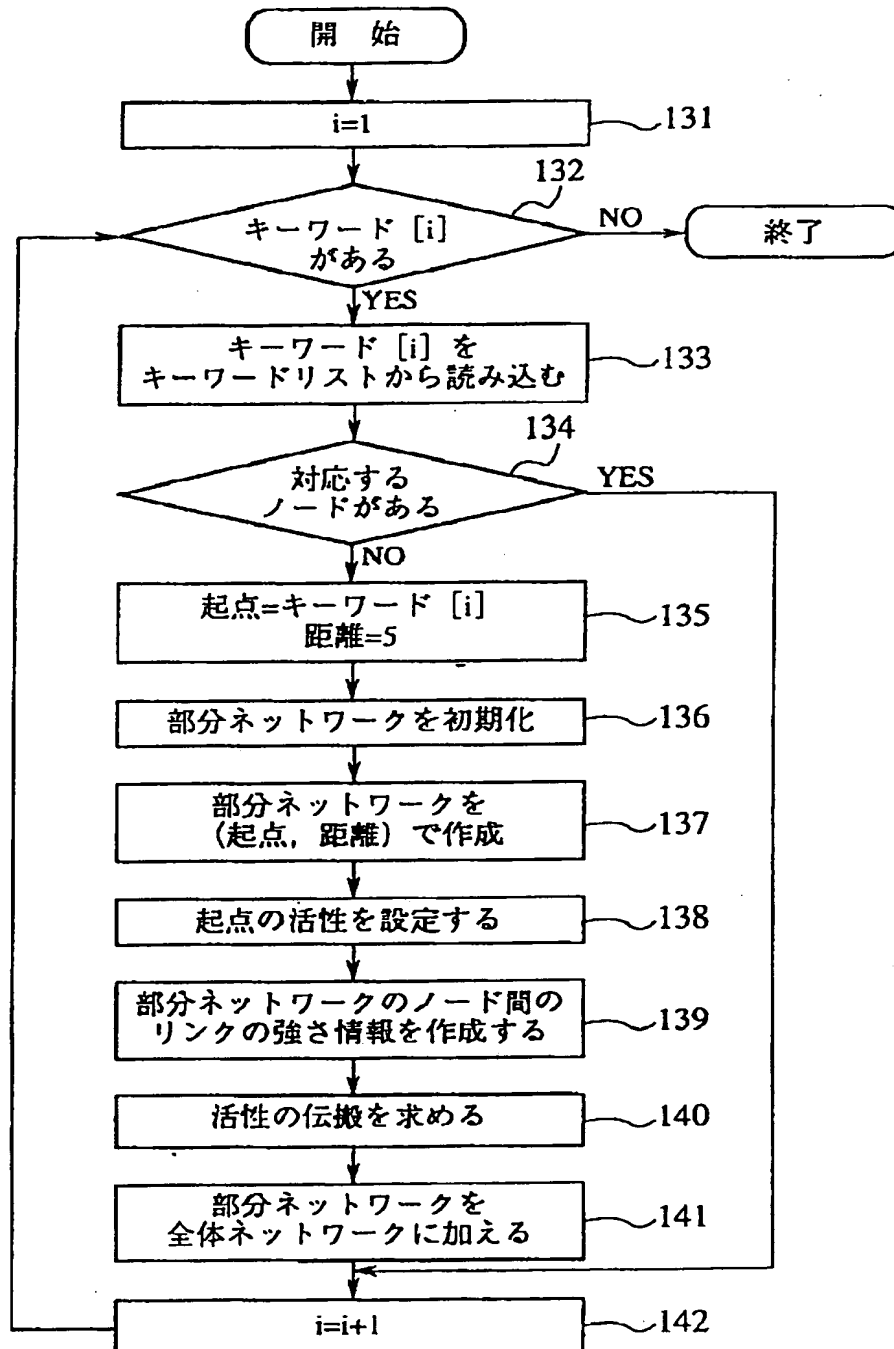
【図10】



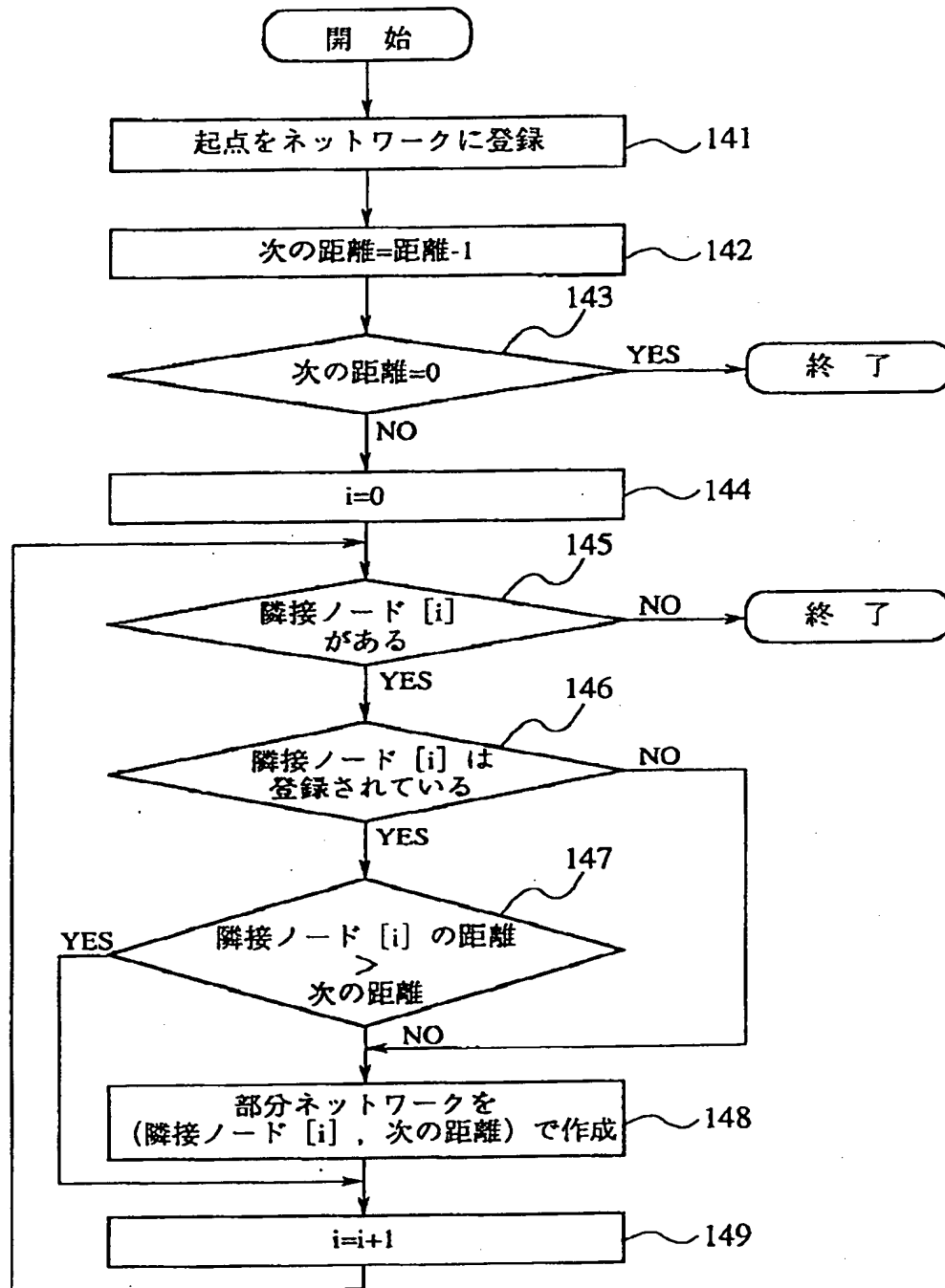
【図11】



【図13】

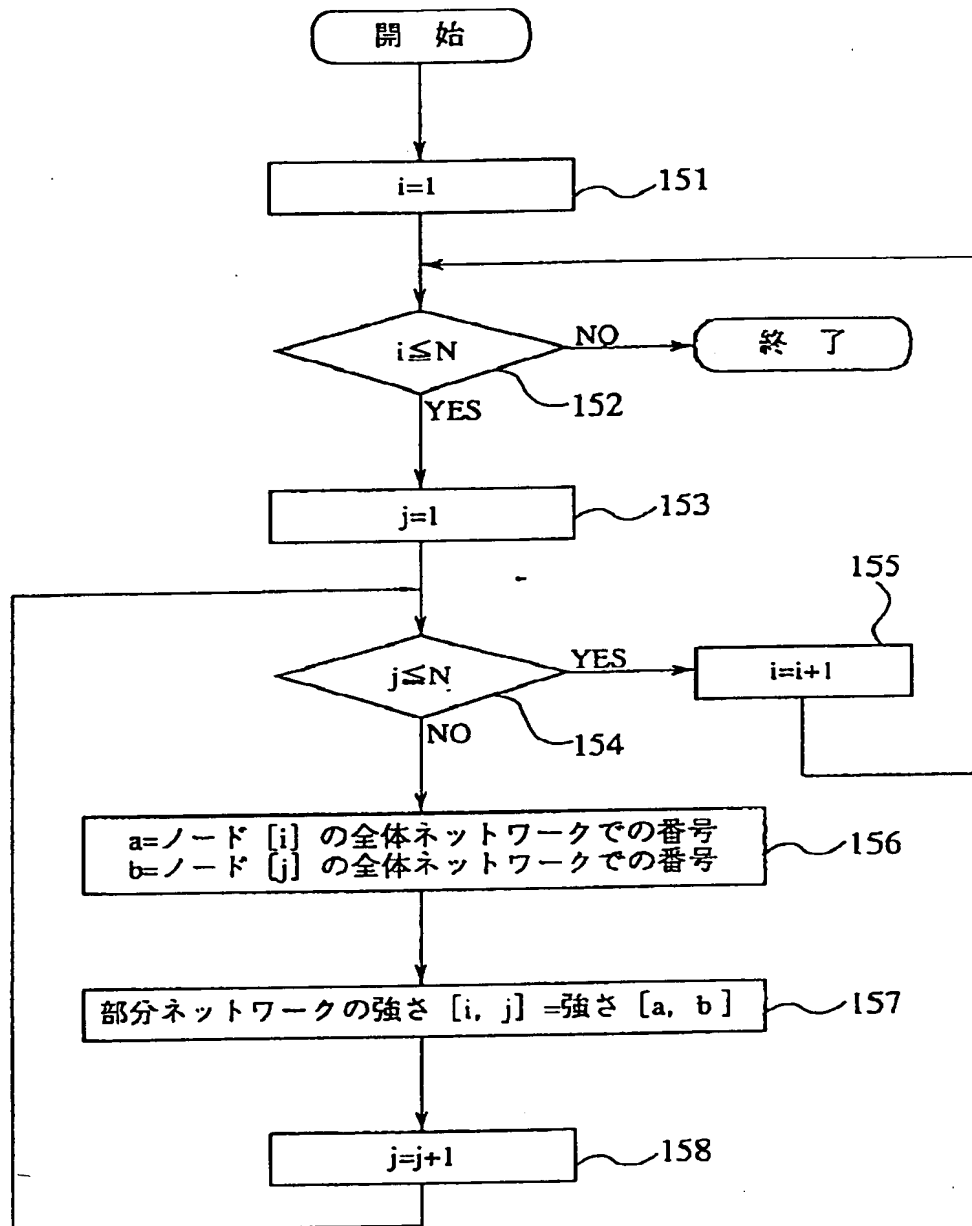


【図14】





【図 15】



【図16】

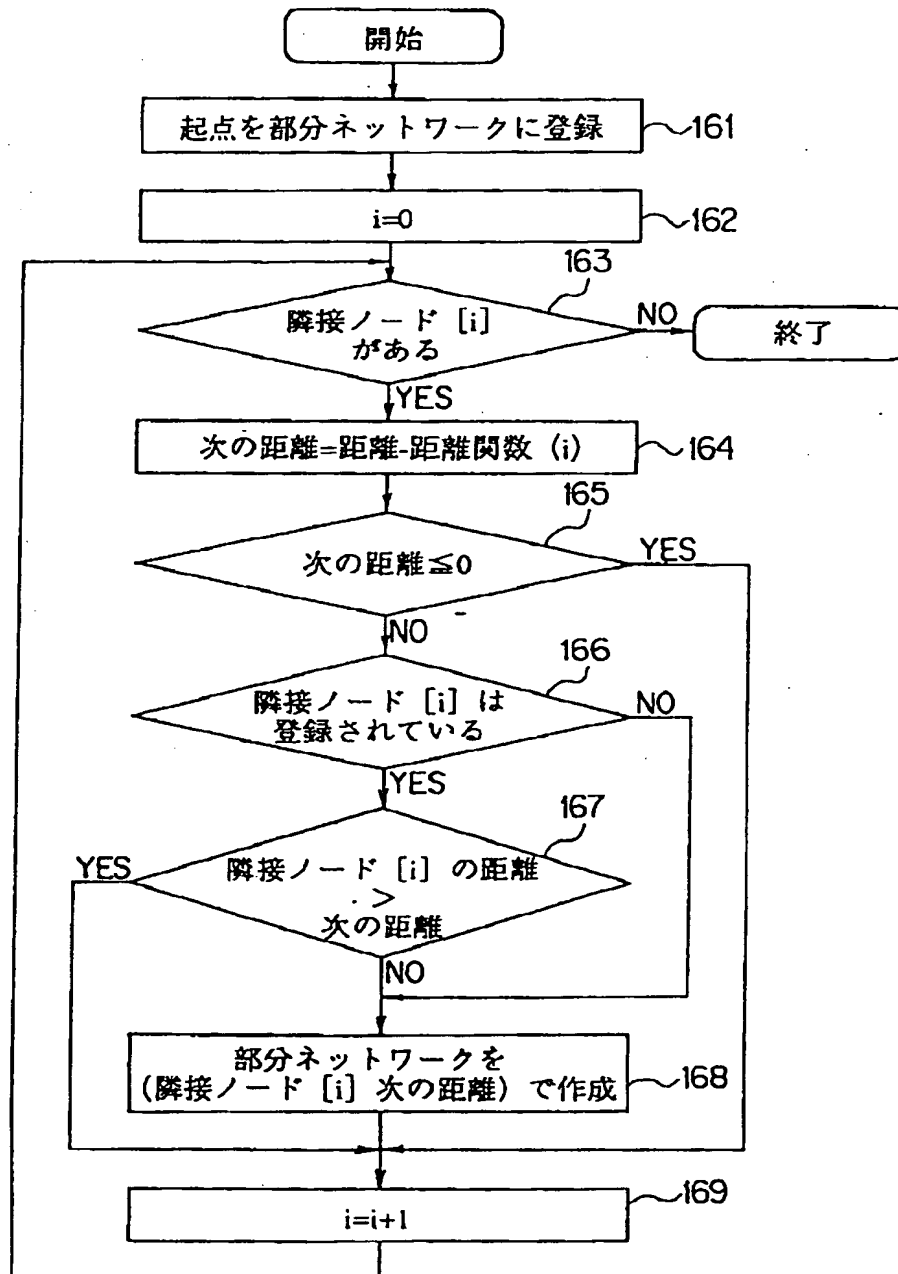
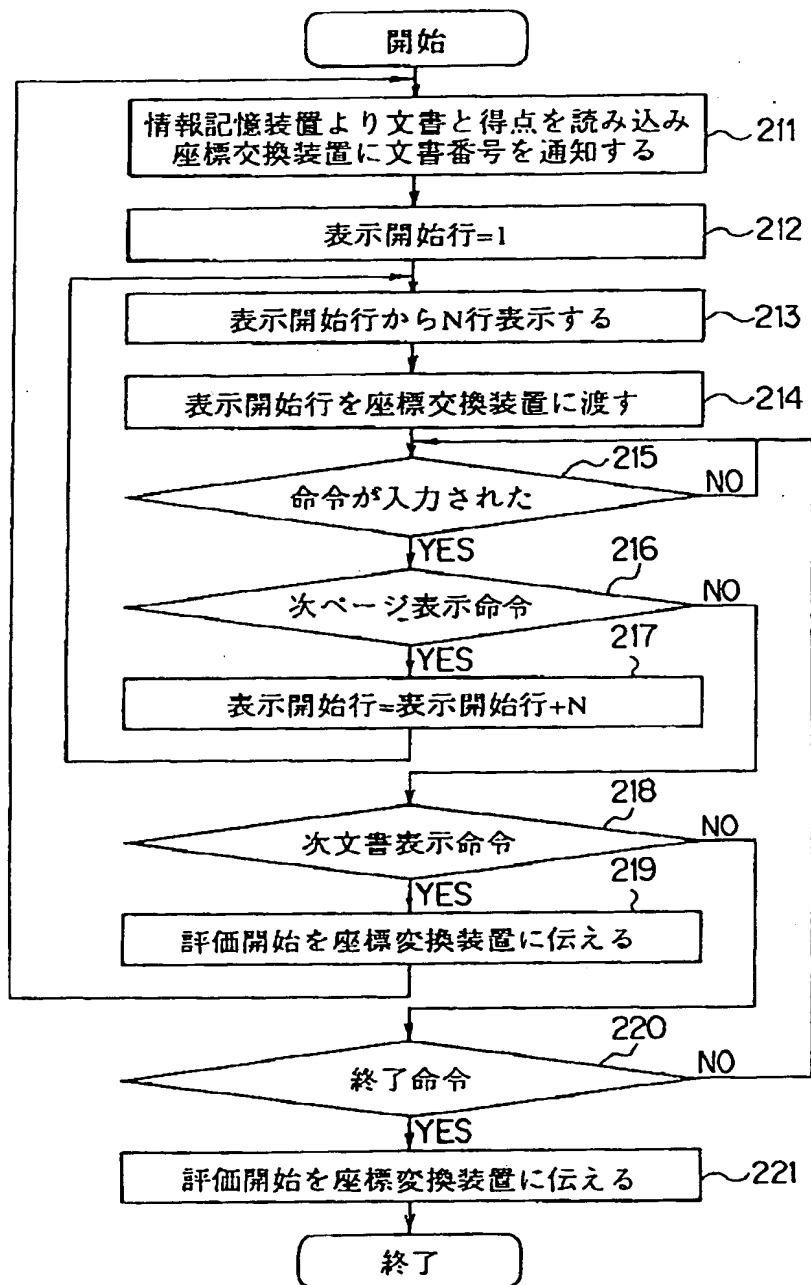


Figure 1 shows a 30x30 grid of Japanese characters (katakana) used for font rendering analysis. The grid is divided into three horizontal sections. The top section (rows 1-10) contains the character 'あ' (a) in the first 10 columns and 'い' (i) in the last 20 columns. The middle section (rows 11-20) contains 'お' (o) in the first 10 columns and 'か' (ka) in the last 20 columns. The bottom section (rows 21-30) contains 'こ' (ko) in the first 10 columns and 'さ' (sa) in the last 20 columns. A black rectangular box highlights the area from row 11 to row 20 and column 11 to column 20. To the right of the grid, there are two labels: '文書全体' (Entire Document) with a line pointing to the top right, and '表示領域' (Display Area) with a line pointing to the highlighted box. Above the grid, the numbers 1 through 30 are listed vertically on the left, and 1 through 30 are listed horizontally at the top. Above the top row of characters, the numbers 1 through 12 are listed horizontally.

Figure 1 shows a 30x30 grid of Japanese characters. The grid is divided into three horizontal sections. The top section (rows 1-10) contains the character 'あ' (a) in the first 10 columns and 'い' (i) in the next 10 columns. The middle section (rows 11-20) contains 'う' (u) in the first 10 columns and 'え' (e) in the next 10 columns. The bottom section (rows 21-30) contains 'お' (o) in the first 10 columns and 'か' (ka) in the next 10 columns. The grid is labeled with '1' to '30' on the left and '1' to '30' on the top. A bracket on the right side of the grid is labeled '表示領域' (display area).

【図20】



【図26】

		1	1	1	1	1	1	1	1	2											
	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	
1	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ
2	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ	あ
3	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い
4	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い	い
5	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う
6	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う	う
7	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え
8	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え	え
9	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お
10	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お	お
11	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か
12	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か	か
13	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き
14	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き	き
15	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く
16	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く	く
17	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け
18	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け	け
19	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ
20	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ	こ
21	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ
22	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ	さ
23	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し
24	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し	し
25	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す
26	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す	す
27	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ
28	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ	せ
29	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ
30	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ	そ

文書全体

表示領域

文書全体

表示領域

【図27】

